THREE MUNDA SCRIPTS

Norman Zide
University of Chicago

1.0. INTRODUCTION

"Tribal" central India in the first half of the twentieth century was a place where many new scripts were devised, devised by native speakers of the languages. Apparently it was felt by the newly conscious speakers of "tribal" languages that a full-fledged language in the Indian context needed a script of its own, one clearly different from those of its neighbors. Certainly more than a dozen of these scripts were made for less than half that many languages. For some of them, e.g., Ho, more than four scripts were devised. Most of these scripts are no longer used, or even remembered. Several of the tribal communities were, sooner or later, satisfied to use a regional or international scripts, and speakers of these languages saw—and were pressured to see—the advantages of learning a regional language, and, in some cases, replacing their native languages with the dominant regional languages: Hindi, Bengali, Oriya, Telugu, Marathi. Most of the Christian missions and the communities influenced by them in this area used the Roman script until fairly recently, when it was replaced by a local regional script. (This supplanting of Roman did not take place in northeast India or the Nicobars.) The Christian missionaries were in the earlier periods the chief—in many regions the only—advocates and promoters of tribal literacy in the tribal languages, and in the regional languages. The new scripts were designed on a variety of principles and in various styles, borrowing from and/or dissimilating themselves from the familiar scripts of their neighbors. Some of the scripts, notably the Ol script which Raghunath Murmu devised for Santali, show independent thinking about sounds, symbols, words and their graphic representation.

The three scripts discussed here—each used for writing a language of the Munda family—were devised by charismatic community leaders as parts of a comprehensive cultural program; in all three cases they were offered as improvements on scripts used by Christian missionary linguists and their "tribal" associates. All the devisers of these scripts, unlike Shong Lue Yang, the creator of a Hmong writing system the origins of which are in other respects

rather similar,[1] were familiar with one or more scripts used in their provinces. These three scripts survive; they are in use in primary and adult education, and are the vehicle for a variety of printed materials. In the case of Santali we have a language spoken over a widespread area by more than six million people.[2] It was—and is—written in four "older" scripts: Devanāgarī, Bengali, Oriya and Roman, the areas being as multiscriptal as "Greater Kurdistan" and some parts of Southeast Asia. Raghunath Murmu's Ol script has been the most successful of the three scripts, and there have been recent attempts by Santals to induce other "tribal" groups in the Chota Nagpur area, both Munda and Dravidian, to adopt this improved script, attempts that, I am told, have met with indifferent success.

Some comparisons can be made among the three, but we lack information on the social anthropology of the uses and meaning of writing—ritual uses among others. On one small topic, controlled comparisons can be made: the preglottalized consonants found in almost all the Munda languages (and many of their Mon-Khmer relatives) are unmarked in word-final position, and alternate with voiced stops in prevocalic position, where they would also be considered unmarked. We can compare how these unmarked pairs of stops are treated graphemically in the three scripts.

2.0. SORA

We know less about Mangei Gomango's Sora script (*Sorang Sompeng,* hereafter "SS") than we do about the other two scripts; the only history and description of this script is found in a short paper by Khageshwar Mahapatra. I have three monolingual booklets in SS, but I cannot make much of them. Sora has been written in a Roman-based script originated by Baptist missionaries, as well as in Telugu characters (devised by G. V. Ramamurti and associates), and, briefly and apparently with little practical use, in Oriya characters also. Mahapatra describes controversy between the promoters of Oriya and those of Telugu for the predominant influence on the Sora people living between the Oriya- and Telugu-speaking populations in what later became the Orissa-Andhra border area. Some "self-conscious tribal leaders," Mahapatra writes, "instead of choosing a side [with which] to merge themselves, endeavoured to maintain their identity by inventing a new script for themselves." Malia Gomango, an influential leader of the non-Christian Sora, led the movement for a separate script, and "inspired his son-in-law, Mangei Gomango" to devise a proper script for Sora. Mangei, "an educated person . . . conversant in Oriya, Telugu

---

[1] See Smalley et al. 1990. [Ed.]
[2] I am not including Assam and Nepal, where Santali laborers were brought in in fairly large numbers.

and English" (and presumably familiar with the Christian Sora script) resigned from his job as a compounder at a pharmacy, and "observed a kind of penance in the hills for several days. Finally, at midnight on June 18, 1936," he received the script "as a divine gift." Mangei said that "the Sora people . . . then . . . were sacrificing cocks, goats, buffaloes and human beings too at their rituals . . . They worship, pay obeisance, and then pour blood and liquor on the deity [deities'] head. However, after many days, at the end of the Age of Kali, the Lord came back and said: Now I have come to you not as Daru [alcoholic liquor] Brahma, but as Akshara Brahma (*akṣara* 'written syllabary character'; the term also has philosophic implications). You worship me in this form. I will be visible on the Məttar Bənom Vijnan hills. The Savara then went and saw. The twenty-four letters appeared in his vision. Then a shrine was built at that site, and the worship of Akshara Brahma commenced from that day." Mangei then set up a new religious order called Məttar Bənom Dəmri, "the religion that opens the eyes and makes people good and wise."[3]

In the Akshara Brahma shrine near Gunupur, in southern Orissa there is an image in the shape of the sacred syllable OM ("written in Oriya characters") which has the twenty-four letters of SS, twelve numerals and a crest. Note the use of OM in the Ho Varang Kshiti script. The mantric power of OM and its peculiar written representation seem more unique than prototypical, but certainly are important enough to be taken over, and ideas of OM influence notions of writing and uttering and what these can effect. What discourses and practices are associated with the various Hindu-influenced notions (Daru Brahma, Akshara Brahma) among the Hinduized Sora of the Gunupur region is not clear to me, nor who "the Lord" is, or what importance a Kali Age has. The various Gods whose names contribute to the letter names of SS are, perhaps, loosely and syncretistically linked with these wider Hindu notions.

One would like to know whether the various less Hinduized, interior Sora groups accept the script and the various discourses and practices that Mangei would like to go with it, and what they make of them. Mahapatra writes that Mangei has proselytized "a good number of tribesmen, and has established sub-centres all over the Sora-speaking tract in Orissa and Andhra Pradesh. In many villages evening schools have been set up to teach the script, and in some areas the script has been widely learned, and is used in intra-community communication and literary activities." It is not clear, however, how many Sora are literate at all, and in what script(s). Probably a fairly small minority. Mahapatra writes that after the first book in SS (by Mangei) was printed in Vijayawada, a press was set up locally, at Dambasara (Putsahi). Mahapatra

---

[3] David Stampe confirms my understanding that the Sora did not sacrifice humans. Why make the statement above, and for whom? Non-Soras?

writes: "the precinct of the press is treated as a place of sanctity," and is the center of the SS literacy drive. Mangei, who wrote all the publications listed by Mahapatra (eight of them, published between 1967 and 1976, after the first one), "mostly lives there as the sire of the institution." Mahapatra notes that the press has also published tracts, almanacs, invitation cards and bulletins. It is not clear whether these genres, however modified, are in any way traditionally Sora, or were taken—translated—from Oriya and Telugu. The SS script seems to be in large part an instrument of acculturation: note the illustrations in the SS booklets of Lakshmi, Nehru, Ganesh and a rather elegant teacup and saucer.

The twenty-four characters are arranged in a four-row-by-six-column diagram, the six vowels being entered in the last (bottom) row. See Table 1. (The schwa vowel is "inherent," i.e., is not represented by a written character.) The twenty-four letters get their names from twenty-four gods in the Sora pantheon, thus (working across the top row) {s} for "Sundaŋ", {t} for "Tənod", etc.[4] Mahapatra found no rationale for the arrangement of the characters and/or of the gods that give them their names. It is not clear how the letter names or the sounds they symbolize relate to the particular gods from which their names are derived. (There is probably no esoteric sound symbolism involved, of the Ho type or any other.) The names of the consonant characters are formed by adding "aʔ" to the consonant sound, i.e., *saʔ, taʔ*, etc. Mahapatra suggests that the characteristic shapes of the SS characters owe something to English cursive letter shapes, which could have been familiar to Mangei; this seems plausible. Perhaps the loops and curlicues were influenced by the Telugu script, but there seems little influence of the regionally dominant Oriya script. No attempt at a graphemic componential analysis of the characters is offered here. We note only that the letters are clearly different componentially from the numbers, which is common elsewhere.

2.1. Sounds and symbols

Divine providence does not guarantee that a script will have linguistic efficiency. Mangei's script does not represent the phonemes of Sora as well as it might. However, it should be noted that we do not know as much about the script as we need to, so that there may be morphophonemic and perhaps

---

[4] I write phonemes and morphophonemes between slashes (e.g., "/e/"), graphemes between curly brackets (e.g., "{o}"), and allophones and phonetic descriptions between square brackets (e.g., "[o]"). Both Sora scripts (like the other non-Perso-Arabic-derived scripts of India) are written from left to right.

(𝑙𝑜2𝑒3𝑦𝑇 𝑜𝑈𝑦)

s	t	b	c	d	g
m	ŋ	ḷ	n	v	p
y	r	h	K	j	ñ
a	e	i	u	o	E	maE

1	2	3	4	5	6	7	8	9	0

Note: The last symbol in the fourth row is not a vowel, but a diacritic. [Ed.]

Table 1. The tabular arrangement of the Sora characters.
(Source: Mahapatra 1978-79.)

dialectal information built into Mangei's writing system that we are missing.

Sora phonology—vowel phonology in particular—is still problematic.[5] The standard analysis of the Serango dialect gives the following inventory:

i	ɨ	u		p	t			c	k	ʔ
e	ə	o		b		ḍ		j	g	
ɛ	a	ɔ		m	n			ñ	ŋ	

		r	ɽ	l	w	

Figure 1

The arrangement of the graphemes is as follows:

s	t	b	c	d	g
m	ŋ	l	n	v	p
y	r	h	k	j	ñ
a	e	i	u	o	ɛ

Figure 2

There is some question about whether the positions of {e} and {ɛ} in Figure 2 should be reversed.

There are problems of loan phonology. The SS writing system makes provision for writing certain sounds of standard Oriya (the local lingua franca, Desia Oriya, is not usually written), e.g., the retroflexes /ḷ/ and /ṇ/, and the various aspirates, voiced and voiceless. Since Sora is phonologically conservative and keeps what is probably the old Munda pattern, dental *t* (but not retroflex *ṭ*) and retroflex *ḍ* (but not dental *d*), the pattern-filling *ṭ* and *d* of Oriya also need representation. The need for representing the non-Sora Oriya sounds seems more hypothetical than actual, since the conservative Sora habit has been to Soraize, i.e., to adapt to Sora phonology, rather than to take in foreign elements. (Ho and Santali have incorporated more of the phonology of their Indo-Aryan neighbors, although the markedness treatment of preglottalized and depreglottalized consonants is handled differently in each of these languages.)

Roman Sora uses no diacritics on consonants (neither for dental 'd' nor retroflex /ṭ/, nor for anything else), and with vowels only the tilde for nasalization; it does not mark aspiration in the few loans that turn up. These are

---

Soraized—e.g., 'Peter' (in Oriya transcribed with a retroflex [t]) is written *Pitor* with the normal dental Sora *t*, 'Matthew' is written *Mattiu*, etc. Roman Sora uses capital letters for proper names, like English. SS has no capital letters. I do not know what sound word-final {rr} represents. Mahapatra writes that the Oriya retroflexes /ʈ/, /ɳ/, /ɭ/, /ʂ/ are written in SS as *tD, nD, lD, sD*, with the *D* here representing the one Sora diacritic, *mae* (Mahapatra calls it an "auxiliary letter").[6] The aspirates are written—if they are written—by adding "Sora j" after the (unaspirated) stop symbol, e.g., {bJ} for Oriya {bh}. The aspirates in the Oriya writing system are written with unitary symbols.

There arises the question of the marking of morpheme-final preglottalized *'b, 'd* and *'j*, and the marking of the (non-preglottalized) voiced homorganic stops in the allomorphs of morphemes with *'b, 'd* and *'j* where these are followed by V. In Roman Sora there is no diacritic or special character representing the preglottalized consonants, and thus no distinctive way of indicating the contrast between the two. Given the distribution of, e.g., ['b] vs. [b], glottalization or non-glottalization of the consonant written would be predictable in most but not all cases. In SS there are no conjunct consonant characters either, and no contrastive representation of preglottalized consonants.

Roman writes the glottal stop with a raised apostrophe (e.g., {ɔ'ɔn} 'child'); SS uses the *h* symbol (there is no aspiration in Sora, unless it be written in loanwords; apparently it is usually not written in loanwords), but seems not to write glottal stop in some of the instances where it is phonemically there. Mahapatra notes that there seems to be no use for {c} and {v} in SS. He notes that the retroflex /ʈ/ is written {rD}.

We can show the match between Sora vowel phonology and SS graphemics and at the same time compare the SS transcription with Roman. Roman Sora has six "basic" vowel graphemes {i, e, a, ə, o u} and two rare ones: {ɔ} and something written with a kind of *u* or omega which I write {U}. The sound represented by {U} seems to be an allophone of /u/, a lower, laxer vowel (in, e.g., {bUñaŋ} 'brother'. The {ɔ} represents certain instances (perhaps phonetically distinct allophones, perhaps allophones further distinguished grammatically—probably both) of what would be transcribed phonemically with /ɔ/, but not all or even most of the instances of phonemic /ɔ/ are written with {ɔ}. What determines {ɔ}-marking is not at all clear; it occurs usually word-initially or word-finally, occasionally morpheme-finally, e.g., in the clause-final interrogative word {pɔ}, and in words where variation or ablaut relations between /a/ and /ə/ are involved. There may be morphophonemic information in this transcription that I am missing. In Roman, {ə} probably systematically represents /ə/ and /ɨ/. In SS, Mahapatra writes, /ɨ/ is

---

6 This diacritic is to be found at the rightmost end of the vowel row in Table 1. [Ed.]

(presumably everywhere) graphemically merged with /i/ and written {i}. (I have some doubts about this.) In Roman, {e} represents /e/ and /ɛ/, while {o} represents /o/ and those instances of /ɔ/ not written with {ɔ}.

SS writes seven vowels: the "inherent" schwa—indicated by writing a consonant character not followed by an overt vowel character (a syllabary property familiar from many Indic scripts)—plus the six vowels written in the fourth (bottom) row of the diagram of characters (Table 1). These are {a, e, i, u, o, ɛ}. Note that the mid vowels /ɛ/ and /e/ are distinguished, but apparently {ɔ} and {o} are not. (There is a possibility that /ə/ and /ɔ/ are wholly or partly merged, graphemically, as "inherent schwa.") SS presumably merges, as Mahapatra notes, /ɨ/ and /i/ as {i}—I suspect some graphemic merger with {ə}). Sora /o/ is graphemically merged, either with /u/ or (more likely) with /ɔ/—and is written either {o} or {u}.

It is not clear how SS writes vowel nasalization.

SS writes geminate {a} (e.g., in {aamn}, i.e., /amən/ 'arrow') where the gemination of {a} seems to carry some morphophonemic (ablaut?) information about the initial vowel, and, perhaps, also some indication of phonetic length/stress. Note that standard Oriya (but not Desia Oriya, the lingua franca in much of the Sora country) has an /o/ (tense, higher mid), contrasting with /ɔ/ and /u/, but nothing paralleling this for the front vowels.

phonemic	/i/	/e/	/ɛ/	/a/	/ə/	/ɨ/	/u/	/o/	/ɔ/
SS	{i}	{e}	{ɛ}	{a}	{ə}[7]	{i}	{u}	{o}	{o}
Roman	{i}	{e}	{e}	{a}	{ə}	{ə?}	{u, U}	{o}	{o,ɔ}

Figure 3

2.2. Consonant clusters and schwa-deletion indicators

There are no conjunct consonant characters (as in the Indic scripts) in Roman or in SS. Consonant clusters are written in Roman with sequences of separate consonant characters. In a system like SS, the inherent schwa should automatically be read between any sequence of consonant characters,[8] so strictly speaking there should be a *halanta*-like diacritic to delete the schwa in cases of real consonant clusters. Failing this, there might be prosodic and/or morphophonemic rules which govern its deletion, so that it need not be indicated in the writing. If no diacritic is used, and if the deletion is not predictable, the fluent reader must simply learn whether a schwa is to be read or

---

7 This vowel is "inherent" in the absence of any other vowel sign.
8 Mahapatra is confusing on this point.

not. There is no *halanta*-like diacritic in SS, so readers are left to their own
devices.

3.0. SANTALI

Santali in the *Ol Cemet'* (OC) or, as it is usually called now, *Ol Ciki* (or just
Ol) script has none of the syllabary properties of the Indic scripts, not even the
minimal schwa-indication arrangement of the Sora SS script, and no
complications in the placing of vowel signs. All the vowels are written with full
letter symbols, written in linear sequence, following the "preceding
consonants," if there are any. The Santali represented in OC is the southern
dialect, spoken in the Mayurbhanj district of Orissa. What is distinctive about
its phonology is that this dialect has six vowels, whereas the dialect of Santal
Parganas (described by Bodding and others) has eight or nine. The literature
from Raghunath's Ol Press says nothing about how the additional vowels are to
be written, but Mahapatra and Mahapatra explain how this is done by adding
diacritics to some of the six basic vowel characters to yield a total of nine.
Pandit Raghunath Murmu, the inventor of OC, arranges the written characters
in a six-row-by-five-column diagram, the six characters in the first (leftmost)
column being the six vowels in the order (from top to bottom) {ɔ a i u e o}. All
the consonant letter names in a particular row have names with the vowel found
at the beginning of that row. Thus the first row has the letters {Lɔ, ɔT, ɔK', ɔŋ
and ɔL} for /ɔ/, /t/, /k'/ (i.e., [k', g]), /ŋ/ and /l/. For Raghunath's diagram
giving the sounds and the letter names, see Table 2. For the letters in printed
and cursive form and their 'objects', the pictures from which the letter shapes
derive, see Table 3. See Table 4 for all five scripts used for Santali as presented
in a leaflet from the Ol Press.

The inventor of the OC script was usually referred to in correspondence
with me (by the representatives of the Adibasi Cultural Association at the Ol
Press in Rairangpur, Mayurbhanj District, Orissa) as "Pandit Raghunat Murmu
D. Litt", the followers of Raghunat(h) and promoters of the script being always
aware of the importance of academic credits and accomplishments in selling the
script to the larger Indian community and thereby to a lesser degree to the
Santals themselves, particularly those outside Raghunath's home territory in
Mayurbhanj.[9] Raghunath devised and advertised his script as easy to learn—
easier than the other four Ol Urum scripts; Ol Cemet' was defined as "easy to
learn" compared with Santali written in the other scripts (Oriya, Bengali

---

[9] This sort of accreditation was not needed, or available, for Shong Lue Yang in the Hmong
world.

Norman Zide

SANTALI

		I	II	III	IV	V
sound	i	ɔ	t	k'~g	ŋ	l
letter name		ɔ	ɔt	ɔk'	ɔŋ	ɔl
sound	ii	a	k	c'~j	m	w,v
letter name		a	ak	ac'	am	aw
sound	iii	i	s	?(~h ?)	ɲ	r
letter name		i	is	i? (?)	iɲ	ir
sound	iv	u	c	t'~d	ṇ	y
letter name		u	uc	ut'	uṇ	uy
sound	v	e	p	ḍ,r	n	r,l
letter name		e	ep	eḍ	en	er
sound	vi	o	ṭ	p'~b	w̃	h
letter name		o	oṭ	op'	[ow̃	oh]

Table 2. The Santali sounds and letter names of
Pandit Raghunath Murmu, D. Litt.

O (ɔ) T (t) K (ɡ) N (ɲ) L (ɛ) U (w) C (c) T (ʈʰ;) R (n) Y (y)

A (e) K (k) C (cʔ) M (m) W (w) E (e) P (p) D (ḍ) N (n) R (ṛ)

I (i) S (s) H (ʔ) N (ɲ) R (r) O (o) T (ṭ) B (pʔ) W (w̃) H (h)

1 2 3 4 5 6 7 8 9 0

*Table 3. Printed and cursive Santali letter forms and the pictures
from which they are derived
(Source: Zide 1967.)*

Norman Zide

᭒	O	ᬇ	᭓	ρ
ᝌ	(ᯮᯓ) ᯭ	(ᯮᯕ') ᯕ,'	(ᯮ') '	(ᯮᯞ)ᯭ
ᝏ	(ᯰᯓ)ᯱ	('ᯕᯓ)ᯕ,'	(ᯰᯛ)ᯬ	(ᯰᯕ)ᯭ
ᝐ	(ᬰᯗ)Ϙ	(ᬰᯕ')ᬰ,'	(ᯕ')˙	(ᬰᯗ)Ϙ,
O	(O'ᯗ)t	(O'k')k'	(O'n˙)n'	(O'l)l
᭓	ᖮ	ᑌ	ᑕ	ᗐ
ᝊ	(ᯮᯗᯕ)ᯕ,	(ᯮᯗᯕ')ᯬ,'	(ᯮᯗᯙ)ᯙ,	(ᯮᯗᯕ)ᯭ,
ᝎ	(ᯰᯗᯕ)ᯱ	('ᯰᯗᯕ')ᯱ,'	(ᯰᯗᯙ)ᯙ,	(ᯰᯗᯕ)ᯭ,
ᝍ	(ᬰᯗ)Ϙ,	(ᬰᯗᗬ)Ϙ'	(ᬰᯗᯙ)ᯙ,	(ᬰᯗᯗ)Ϙ,
ᴧ	(Ak)k	(Ac')c'	(Am)m	(Aw)w
ᑫ	Ꮶ	ᗯ	Ϙ	ᗑ
ᯓ	(ᯕᯙ)ᯱ	(ᯱᯕ')ᯕ'	(ᯱᯕ)ᯬ	(ᯱᯭ)ᯭ,
ᯖ	(ᯕᯙ)ᯙ,	(ᯱᯕ')ᯕ'	(ᯱᬶᯬ)ᯬ,	(ᯙᯙ)ᯙ,
Ϙ	(ᬰᬰ)Ϙ,	(ᬰᬰ')Ϙ'	(ᬰᯟ)ᯟ	(ᬰᬰ)Ϙ,
ᒋ	(1s)s	(1h')h'	(1n')n'	(1r)r
ᑕ	ᗷ	ᑤ	ᑬ	ᗌ
ᯕ	(ᯰᯟ)ᯥ	(ᯰᯗ')ᯗ,'	(ᯰᯞ)ᯞ,	(ᯰᯙ)ᯙ,
ᯖ	(ᬰᯗ)ᯥ,	(ᬰᯗ')ᯗ,'	(ᯰᯙ)ᯙ,	(ᬰᯗ)ᯙ,
Ϙ	(ᬰᬰ)Ϙ,	(ᬰᬰ')Ϙ,'	(ᬰᯗᯟ)ᯟ	(ᬰᯗᯙ)Ϙ,
ᑌ	(Uc)c	(U')t'	(Ur.~)r.~	(Uy)y
ᒿ	ᗄ	ᑏ	ᗃ	ᒪ
ᯕ	.(ᯰᯗ')ᯗ,	(ᯰᗯ.)ᯥ	(ᯰᯟ)ᯟ,	(ᯰᯥ.)ᯥ,
ᯗ	(ᯰᯕ)ᯞ,	(ᯰᯕ.)ᯗ,	(ᬰᯙ)ᯙ,	(ᯰᯥ.)ᯥ,
᭍	(ᯰᬱ)Ϙ,	(ᯰᯗ.)Ϙ	(ᯰᯗ.)Ϙ,	(ᯰᬰ)Ϙ.
E	(EP)p	(Ed)d	(En)n	(Er.)r.
᭓	ᑎ	ᑗ	ᑬ	ᗒ
᯵	(ᯮ᯲ᯗ)ᯥ.	(ᯮ᯲ᯕ')ᯗ,'	(ᯮ᯲ᯥ)ᯥ	(ᯮ᯲ᯥ)ᯥ,
᭍	(ᯰᯗ.)ᯥ,	(ᯰᯗ.')ᯗ,'	(ᯰᯥ.)ᯥ,	(ᯰᯥ.)ᯥ,
᭍	.(ᬰᯗ.)ᯥ,	(ᬰᯗ.')Ϙ'	(ᬰᯥ.)Ϙ	(ᬰᯥ)Ϙ:
O	(Ot.) t.	(Op')p'	(OWn.)n'	.(Oh)h

The OC letters are on top and beneath them Hindi,
then Bengali, then Oriya and lastly English, i.e. Roman.

Table 4. The five scripts of Santali
(*Source:* Adiabasi Cultural Association, ca. 1957.)

Devanāgarī and Roman), which Raghunath called *Ol Urum*, i.e., 'dusty (superannuated) writing'.

Among the features making it easier to learn, the following were mentioned:

(1)   The script made use of signs and symbols long familiar to the Santals, e.g., marks made on stones or trees to indicate 'danger', or 'meeting(place)'.

(2)   The "scientific" arrangement of the letters "facilitates the children to learn the name of letters as there is a flow being headed by the same vowel," i.e., the property described and exemplified above with the ɔ-row.

Although Raghunath was familiar with the Sanskrit-derived diagrams (*varṇa samāmnāya*), the phonetic properties organizing these, and the separation of vowels from consonants, do not figure in Raghunath's system.

Some of these properties are exemplified, irregularly, in the English and other European alphabets: an initial vowel at the beginning of each 'set' or 'row' of letters, but not, e.g., *ay, bay, say, day*, etc., although there are sporadic patches of same-vowel names in a set. Thus, English letters can be— and were in teaching materials used in India?—ranged as follows:

A	B	C	D		
E	F	G	H		
I	J	K	L	M	N
O	P	Q	R	S	T
U	V	W	X	Y	Z

In English too one can find nouns (*bee, jay, cue*), pronouns (*I, you*), verbs (*see, pee, are*) and interjections (*oh*) in the letter names, but these meanings are taken as accidental and irrelevant, and are not reflected in their letter shapes. The OC letter names are all open syllables: those of the vowels of CV shape, those of the consonants of VC shape. The English letter names are not so neat.

(3)   The shapes of the letters are not arbitrary, but reflect the names of the letters, which are words (usually the names of objects) represented in conventionalized form in the pictorial shapes of the characters. These "objects" include some actions, i.e., verbs, as well as nouns and pronouns, e.g., *to blow, to write, to harvest, to thresh grain, to avoid, to point out a place*. The vowels /ɔ a i u e/ (for /o/ see below) are "modified from pictures" of actions, 'burning' for {ɔ} (letter name Lɔ), 'digging' for {a} (LA), 'bending' for {i} (LI), 'lifting water by dipping a ladle' for {u} (LU), and 'swelling' for {e} (LE). Interjections in letter names include ɔK', presumably the sound

indicating vomiting, and *I?* 'a word expressing terror'. "It is modified from a picture of an action of a man in terror. It is also a symbol of danger."

The vowel /o/ is described as "a low pitch on ɔ, and is modified from a picture of mouth pronouncing [o]." What Raghunath means by "low pitch" is unclear. The phone [o] is usually described as a back vowel, higher than [ɔ], with no distinction in pitch, as that term is commonly used in linguistics. It is conceivable that there is a distinction in phonation type—perhaps having to do with laryngeal tension—since both Santali and Mundari are distinctive in this regard, but this seems unlikely. (Note that the Sora SS transcription also suggests possible phonetic distinctions not recorded by the linguists who described the language previously.) Thus the letter {ḍ}, called EḌ, means 'to point out a place'. It was modified from a picture of a fist with a finger pointing to a particular place. It was previously a symbol for 'meeting place'.

Similarly, the letter {n}, called EN, means 'to thresh grain'. It is modified from a picture of two legs threshing grain; {ṛ, l, ḷ (?)}, called EṚ, means 'to avoid', and is modified from a picture of a path turning to avoid an obstruction or a dangerous place. It was also a previously used symbol; {c'}, called AC' "means 'he' [or 'she']." "It is modified from a thin picture of a man with his left hand pointing to a man (3rd person)." OH and Õw are not pictures of anything.

How useful these are in learning a script I cannot say, but the claims of a connection with earlier writing/marking, relating the shapes of the characters to everyday objects and activities, were on the whole attractions and features that the other (Ol Urum) scripts lacked—and could be incentives to learn it.

For three of the characters the shape represents the shape of the lips when pronouncing the letter name, thus /ŋ/ (Oŋ) "means 'to blow air'". This letter is modified from a rough sketch of upper and lower lips with air coming out between the two; /o/ (LO) is "modified from a picture of [the] mouth pronouncing O (low pitch on O)"; and {k'} "is modified from a rough sketch of opened mouth to vomit when [ok'] is pronounced."

One ingenious—and scientific—feature that certainly increases the efficiency of writing Santali, and one unique to OC, is the deglottalizing OHOD (ɔhɔt') diacritic, which "is a signification to open the checked consonants to simple consonants and used only with checked consonants." This neatly preserves the morphophonemic relationships between the "checked consonant" and its homorganic voiced equivalent, the former occurring word finally and at certain word-internal preconsonantal junctures, and the latter occurring prevocalically but never morpheme-initially in these alternations. In the material I have, it is the glottalized (checked) consonant that is the unmarked one. Thus,

OK' is the name of a letter that represents both [k'] and [g]; there is another letter for /k/ but no other for /g/. [gh] would be written with {OK'} and the aspiration diacritic. The diacritic for aspiration (OH) represents /h/ and the aspiration of voiced and voiceless consonants. It is included in the letter diagram, as is the rare Ow̃ (also in the sixth row), described as a "nasalized [w]." Another diacritic is the raised dot for vowel nasalization; as Mahapatra and Mahapatra point out, Raghunath learned something from Bodding's work on Santali and the analyses it was based on, and the representation of nasalization is probably one of the things he borrowed. A horizontal loop added at the top right of the character marks aspiration. This is graphically a diacritic, i.e., something added to the central character, when used to mark aspiration (of a non-nasal consonant), but when representing /h/ the full (non-diacritic) form of the character is used.

Santali, unlike the other North Munda languages, loses certain final vowels (the loss according to Zide and Munda [pers. comm.] is conditioned by laryngeal vowel[s] in the morpheme), and this results in contrasted word-final stops, plain (unvoiced) vs. pre-glottalized, where no such contrast existed in the proto-language (neither in NM [Proto-North Munda] nor in PM [Proto-Munda]), or in non-borrowed vocabulary in the other modern North Munda languages. There was only the one (pre-glottalized) series. Thus, Santali ɔt 'earth, world' (NM *ɔtɛ), the name of the character for /t/, is not to be confused with ut', the name for {t', d}. The final voiced stops, found in loanwords, are written with the pre-glottalized consonant character and the deglottalization diacritic.

Another ingenious, common, and scientific feature is the simplification of character shapes. This of course distances them from the pictorial source representations of their "objects", so that one can describe most of the characters in terms of a small set of graphemic components and writing operations to combine these into characters. (Not all the characters fit neatly— or at all—into this componential analysis.) Thus, one finds characters related to each other in physical shape as are, say, *b*, *d*, *p*, and *q* in written English. Whether this was a desideratum of Raghunath's in deciding on the letter shapes, or recognized as an accomplishment by him and his associates, I do not know. Perhaps not. Most of the characters can be seen as contained in a more or less oval envelope, the letter shape tracing some part of this envelope. (See Zide 1968 for details.) The numerals (as in SS) use another set of components, not those of the letters.

When, thirty years ago, I asked knowledgeable people in Bihar and Orissa what they thought of the chances of a wide—if not pan-Santal—acceptance of

OC, almost all of them were skeptical. The competition of Oriya, Devanāgarī, etc., was too powerful. But, in fact, OC has become more and more widely accepted. Official recognitions have been bestowed. Pandit Raghunath Murmu has been honored by the Orissa government. More recently, various Santal organizations have tried to promote the script for other languages of Chota Nagpur, mostly Munda languages, but for the Dravidian Kuḍux as well. Without much success. The Bengali novelist and social activist Mahashveta Devi, who spent a good deal of time in Chota Nagpur and wrote about the Chota Nagpur tribal scene, has commented (personal communication) that the Santals see themselves as the Brahmans of tribal Chota Nagpur: they believe they have or should have intellectual authority over all the (tribal) communities in the region.

Mahapatra and Mahapatra present much interesting material about Santal script politics in Bengal and Orissa: the official considerations and debates regarding the recognition of one (or more than one) Santali script, the movements advocating political autonomy (e.g., the Jharkhand Movement), and the role of Santal language and script in tribal politics. They describe Raghunath's OC pedagogy in the elementary reading materials he prepared, and point out the extent of the cultural enlargement of Raghunath's programs. These were much richer, more thoroughgoing and thoughtful than Mangei's for the Sora. Thus, Pandit Raghunath Murmu, a poet himself, wrote "a very long epic-heroic play, Kherwar Bir, which serves as a Santal equivalent of [the Hindu Sanskrit epic with versions in the vernacular languages] Mahābhārata."

4.0. HO

Ho is a North Munda language of the Kherwarian branch. It is fairly closely related to Santali. It is less conservative in its vowel inventory than Santali, and has the areally standard five vowel system (but characteristically without vowel length, like most of the other Munda languages with five vowel systems (Korku, Mundari, Kharia, Gutob, and Gorum), and unlike Indo-Aryan or Dravidian). It was provided with a poor, little-used Roman transcription by Lionel Burrows. Burrows was aware of the glottal stop, but saw no reason to write it. He claimed that Ho had no more than four hundred words. Father John Deeney, in the introduction to his dictionary, remarked that he had included twelve thousand items, and there was no reason to think his lexicon was complete. Deeney's (and his Ho associates') Devanāgarī transcription of Ho has been regularly used, particularly by the Catholic Ho community, and fits the language well. Devanāgarī lacks contrastive long and short /e/ and /o/— in Sanskrit and Hindi e and o are necessarily long; otherwise there is no need to provide other vowel symbols or diacritic distinctions, as there would be in the

phonologically conservative Munda languages Santali, Sora and Gta?. Allophonically complex /a/ and (geminate) /aa/ in the North Munda languages do not match or easily map onto the "'long and short a" representations in the Devanāgarī script, and Hindi usage and patterning rather than Ho patterning wins out in determining how Ho is to be written in Devanāgarī. In the glossary at the end of Deeney's Ho grammar the entries are given in Roman, but he does distinguish (sometimes) between (phonetically monosyllabic) long {a} and geminate {aa}. He does this in the Devanāgarī transcription with a raised diacritic like an acute accent. Thus, {hō} 'Ho' should be /hoo/, {hām} 'old man' should be /haam/, etc. A finer-grained analysis of the morphophonology is called for, especially in dealing with long *e* and *o*. In the introduction to his dictionary Deeney says that for the vowels where there is no long vowel symbol in Devanāgarī (i.e., /ē/ and /ō/) he marks length by "putting after the vowel a short stroke slanting up to the right."

In some of the published Varang Kshiti (VK) materials with accompanying Hindi translations, the Hindi form for 'Ho' is properly written /ho-o/ with geminate /o/, whereas in VK Ho the form is written {ho}, perhaps on the tacit assumption that Ho readers know how the form should be pronounced, so that nuances of length and syllabification need not be spelled out, whereas Hindi-based readers need clearer instruction.[10] The VK script developed by Lako Bodra (the script to be described here) is a script for Ho Hayam—very roughly "hieratic Ho"— and the key question here has to do with the difference between Ho Hayam (HoH) and "ordinary Ho," Ho Kaji (HoK). The amount of overlap between HoH and HoK is unclear.

The authority on the matter—and on all aspects of Lako Bodra's script—is H.-J. Pinnow. He suggests three possibilities as to what HoH may be: (a) a (regional?) dialect of Ho; (b) an arbitrary mixture (*willkürlich Mischung*) of Ho, Mundari, Santali and other elements; or (c) an old sacerdotal language (*Priestersprache*) kept secret until now. These alternatives are not mutually exclusive. Pinnow finds the status of HoH to be unclear: "It may be an old secret language of the Ho shamans," he writes. He does call it a distinct dialect, and sees archaic features in it (see the discussion of {HB} below). Certain HoH forms (e.g., /haṛam/ 'old man; husband') are cognate with HoK *haam*, the latter having undergone the Ho rule of loss of intervocalic /ṛ/; but some of the distinctive HoH forms (see the discussion of the numerals) do not seem to be archaic and are not paralleled in any of the other Munda languages or, as far as we can tell, in Mon-Khmer. (A similar situation, in which the differing forms are sometimes archaic, but frequently not, is found in Mundari song forms.)

---

10 See below on the Sanskritizing tendencies of VK for the Ho appropriations of {ṣ} and the Sanskrit {ṛ} vowel.

How old is this sacerdotal language, and how was it formed? Comparative work should clarify the status of HoH as to the archaism of its forms, but comparative work on HoH has been very limited because we do not have a HoH lexicon. (We refer below to Lako Bodra's trilingual HoH manuscript dictionary, but that lexicon was not available to Pinnow, and needs an adept to interpret it to a linguist; one has not yet been found.) For the most part HoH has the same sound system as (Deeney's) HoK: the {HB} is exceptional, and the differences in graphic system reflect different notions of what a writing system for Ho—some dialect of Ho—should have.

Pinnow writes that the script exhibits highly syncretistic features, and many borrowings from a variety of older scripts (I suspect that he overestimates the degree of borrowing), including clear borrowings from the Brahmi script. He sees fourteen of the Ho characters as being fairly directly borrowed from Brahmi. Several of the letter shapes (the non-cursive forms of the letters in particular) resemble English. They are *not*—and this is no accident—similar to Devanāgarī characters. Lako Bodra was not unaware of the Brahmi similarities. According to Pinnow, Lako Bodra's claim was that the script was "invented in the 13th century by a certain Dhawan Turi and rediscovered in a shamanistic vision and modernized by Bodra himself." My own conversations with followers of Lako Bodra in the late 1970s brought out assertions of an even greater antiquity for the script, and the claim that the Ho script was the most ancient in the subcontinent and the (only?) survivor of an ancient flood. Furthermore, the ancient Ho were familiar with DNA, the cuneiform and Harappan scripts, etc. There is an Institute of Ancient Culture and Science Society (*ādi saṃskrti evaṃ vijñān saṃsthān*, but with no Ho name in the booklet I saw) concerned with promoting the Varang Kshiti script and related matters. Ho—any variety of Ho, presumably—must be written in its proper script, i.e., VK, according to Bodra's followers. (In Bodra's earliest views was HoH to be written at all?) The similarities between VK and other Indian scripts are the result of borrowing *from* Ho, not by borrowing *into* Ho, according to Bodra's followers and probably Lako Bodra himself, but according to outside observers the script is his in the sense that he invented it, not rediscovered it. I was told that there are ancient inscriptions in caves and other places, but Pinnow does not record these and I have not seen them or pictures of them.

The greatest interest of VK for linguists is in its social contexts, the claims made for its great antiquity, and how these are supported and used, and Bodra's ideas about writing (see the introduction, "*do shabda*," to his *Ho Hayam Paham Puti* [in Ho and Hindi], and the esoteric (tantra-influenced) ideas of the relations

obtaining between a sound, its production and its written representation.[11] Lako Bodra's manuscript dictionary of Ho Hayam, *Ho Halang Galang*, was commented on very briefly by the present author, but I was not able to penetrate very far into the formidably esoteric apparatus of the manuscript. Bodra connects the shape of each character with an organ of the body, and presumably this organ and its involvement in the production of the sound represented by that character is relevant to the semantics of the words 'containing' these characters, but how this is done—do the semantic dimensions pattern in any systematic way?—is a mystery. The polysemic entries in this dictionary are arranged by semantic registers (this seems to owe something to the format of certain Sanskrit lexicons), e.g., "biological science," "physical science," "philosophy," "technology." Here is one example of his 'definition' of a character, here the character for /n/:[12] "The twentieth *varang kshiti* script (i.e., letter), tenth consonant and fourth nasal consonant . . . It is to be pronounced by pressing the breath amidst the disk [*sic*] to throw out (the sound) through the mouth with buzzing echo (release) through the nose. Being a divine wind the phonetic effects from the testicle to the prostate all over the masculine gender, i.e. the special organ of the male. (The glosses are) in biological science: youth, spirit, ape, gender or sexual part of male, semen, valour, vigour. (Biological science: botanical) the pollen or farina of a flower. Physical science: earth, brick, mortar, kneaded clay. Philosophy: creation, the universe, the world, history (description). Technology: hammer, G.n.p.m. (??) Vishwakarma, artist of the gods, God, mason, carpenter, n. com. (common noun??), egg, semen or vigour, metal, mineral, ore, the root of a verb, wealth."

4.1. Consonants

VK is written from left to right, and the vowel characters are written linearly after the preceding consonants, as in SS and the Ol script. The complications of vowel placement and variant forms initially/postvocalically to be found elsewhere in Devanāgarī, Bengali and Oriya, etc., are eschewed. The characters of VK are now given in a standard tabular form (see Table 5), the consonants being ranged in sets (rows) of three, the first five of these triads having systematic phonetic properties of a kind familiar from the principles underlying the standard arrangement of the Devanāgarī characters. The Ho VK diagram has three columns, not the five required for the Indo-Aryan languages, since Ho lacks the voiced and voiceless aspirate characters. In the Ho chart the order is

---

[11] The title page of this work is given as Figure 1.
[12] The English of Lako Bodra, or more likely of the person who helped him with the English translation, was not adequate to his purposes and is often distractingly bizarre.

218   *Norman Zide*
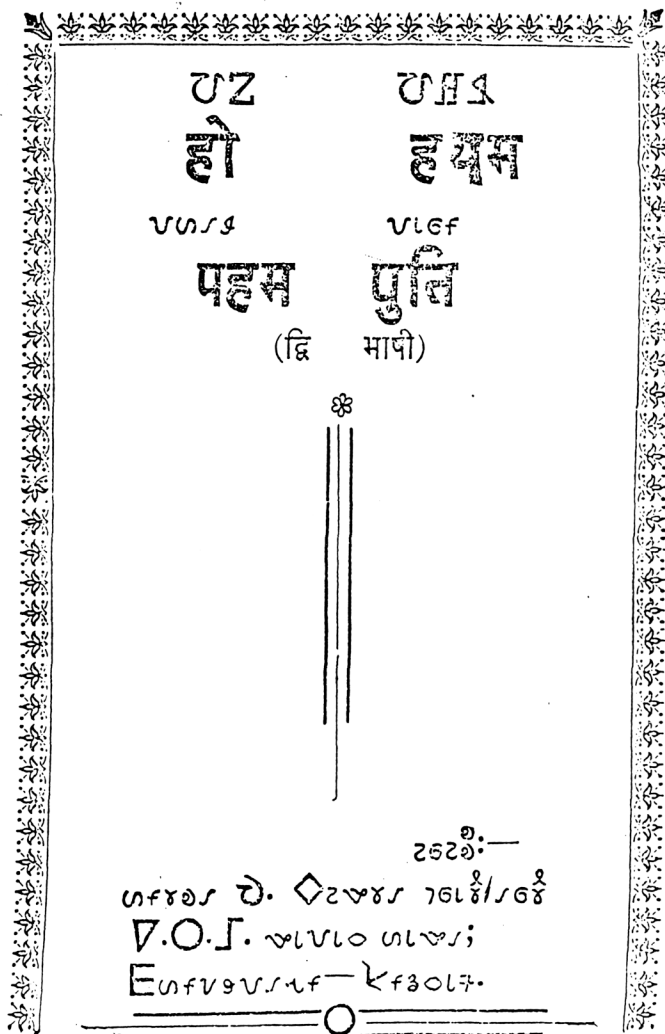
Figure I. Title page of Ho Hayam Paham Puti

Vowels

V	ν	ʃ	ʃ	ⅎ	ꜰ	L	ι	∇	ʋ	Z	ꙅ
Y	ɣ	ꓧ	ß	E	ꜰ	┝	�				

Consonants

ぅ	ȝ	П	ɹ	ዓ	ꝯ
ⴙ	♂	E	ε	ꝯ	ꝯ
⋈	ᴎ	ꜰ	ɾ	L	ι
㇟	⌐	╘	⌐	O	�6
⅃	᧒	◇	◇	U	ʋ
Ʊ	ᴧ	ᄘ	ᴀ	5	ꜱ
Т	ɣ	Ⴂ	ᵧ	⼽	ᵫ

IACS : 2 : 76--77 : 25 : 78-.79 : 500

Table 5. Tabular arrangement of the Varang-Kshiti script for Ho.
(*Source:* Prepared by Lako Bodra's group, probably in 1976-77; also found on
the back cover of a "Memorandum of (the) Institute of Ancient Culture and
Science Society", 1976.)

nasal, voiceless stop, voiced stop (e.g., *m b p*), where the Devanāgarī has the reverse order (*p ph b bh m*). Two other tables, one from Pinnow (Table 6) and one from an intermediate reader prepared at Ranchi University (Table 7), give the phonetic values and transcriptional equivalents in Roman and Devanāgarī. Pinnow is particularly helpful in interpreting Table 7, but the two tables are not in complete agreement, e.g., about the names of the characters (despite what are very likely typographical errors in the Reader). The consonants are the following (in the VK chart order):

ŋ	g	k
ñ	j	c
ṇ	ḍ	ṭ
n	d	t
m	b	p
h	l	ṛ
r	ś	s

The last two rows are miscellaneous, i.e., what's left over (again largely on the Devanāgarī pattern). The retroflex sibilant (ṣ) is not found in Deeney's Devanāgarī transcription; he finds only one sibilant, *s*, and it seems very likely that the phonology of HoH does not differ in this regard from that of HoK. Lako Bodra recognizes an allophonic distinction graphemically, apparently because he wants an archaic, Sanskrit-like cast to VK. *The Sanskritic register of Hindi has* ṣ (the ṣ is found only in *tatsama* or 'unmodified' Sanskrit borrowings), although it is pronounced no differently from ś, which is the common 'sh' character and also occurs in Sanskrit. The retroflex nasal ṇ is more rooted in Ho phonology than is ṣ.

The Sanskritizing can also be seen in the use of the Sanskrit ṛ-vowel { } (pronounced /ri/ in Hindi) to represent /ri/ in the number 'two', /riya/, instead of the usual {ri}, or "*r* plus *short i*". Note also the character names which are written with "archaic Sanskrit-style" initial conjunct consonant characters, but, as one might expect, are not pronounced as consonant clusters; they are pronounced as CVC, e.g., not /hl/ but /hol/, not /hy/ but /hiy/, etc. (see Pinnow). This is not surprising, since non-standard Hindi speakers as well as most Ho speakers would have trouble pronouncing /hl/, /hy/, etc. The letter {h} (see Pinnow's table, Table 6) also, after a vowel, indicates "vowel lengthening" (*Vokaldehnung*), which unexpectedly reminds us of "low tone/aspiration" in Korku (e.g., the reduplication of the verb *ira?*- 'to return' is *i-ì-ra?*, interpreted as *iHra?*). A closer look at the phonetics and the morphophonology of -*Vh*- would be useful.

𝒰ₒ𝒓𝒳ℨ 𝒴𝐹 𝐺 𝐹 *Hbar(a)ŋ citi* „Alphabet"

Originalschr. Groß- Klein- Buchstabe		Transliteration Devanāgarī Latein- Schrift		Bezeichnung der Grapheme Devanāgarī Phonetisch		Phonement- sprechung

1. 𝐸𝑣𝑡 𝑜𝑧𝒳ℨ *ip(a)n bor(o)ŋ* „(einfache) Vokale"

—	—	—	(a), (e), (o)	—	("inhärierend")	/a/, /e/, /o/
V	𝑣	◌	M	m	ङ̇	ŋā·ʔ	/ŋ/, /ŋ/
𝐽	𝒔	अ, आ	A	a	अ̣	ʔa	/a/
𝑓	𝑓	इ	I	i	बि ꞉ इ̣	ʷᵢᵢʔⁱ, ʷᵢʔⁱ	/i/
L	𝑙	उ	U	u	यु꞉ उ̣	ʸᵤʔᵘ	/u/

2. 𝒦ᵥᵟ 𝑗ᵥ𝑡 𝑜𝑧𝒳ℨ *sell ip(a)n bor(o)ŋ* „(gemischte) Vokale"

∇	∇	Ṿ	E	e	�匚Ṿ	ʔεʔ	/e/
Z	𝑧	ओ	O	o	ओ꞉ अ̣	ʔɔʔ	/o/

3. 𝖴𝑧ₒ𝟧𝐹 𝑜𝑧𝒳ℨ *doḅri bor(o)ŋ* „Ligaturen"

Y	𝑦	:	?	?	य꞉	ˈyaʔ	/ʔ/
𝐻	𝒴	य	Y	y	यो	ˈyɔʔ	/y/, /e/, /i/
𝐸	�7	इ,ई,ऐ	I	i	ई̣	ʔi꞉	/i/, /ai/, /ii/
ᖾ	7	उ,ऊ,औ	U	ü	ऊ	ʔu꞉	/u/, /au/, /uu/

4. ◇ᵥ𝟛𝑡 𝑜𝑧𝒳ℨ *beŋn bor(o)ŋ* „Konsonanten"

ठ	ℨ	ङ॰	Ṅ	ŋ	ओङ̣	ʔaŋ	/ŋ/
⋔	𝒔𝑓	ग	G	g	ग॒꞉	gaʔ	/g/
𝐹	𝒮	क	K	k	क꞉	kɔʔ	/k/
ℭ	𝑐𝑏	ञ	Ñ	ñ	विञ̣	ʔεñ	/ñ/

Table 6. The Ho alphabet: the characters with information on letter names, phonetic values, and Devanāgarī equivalents (Source: Pinnow 1972:827-88.)

222   Norman Zide

(Table 6, continued)

Originalschr. Groß- Klein-Buchstabe		Transliteration Devanāgarī Latein-Schrift		Bezeichnung der Grapheme Devanāgarī Phonetisch		Phonement-sprechung
E	ε	ज	J j	विज्न	yü'j	/j/
५	५	च	C c	विच्च	ʔüc	/c/
℧	ᷓ	रग	N ʋ	ऐरग	ʔεʒ̃	/n/
ℾ	ℾ	ड	D ḍ	ओड	ʔɔ'ḍ	/d/
L	ℓ	ट	T ṭ	टॆ:ए	ṭεʔ	/ṭ/
↑	ʋ	न	N n	नुड्ः	nuŋ	/n/
℧	∾	द्	D d	द्:S	daʔ	/d/
O	ϭ	त	T t	ओत	ʔat	/t/
⚐	ℐ	म	M m	Sछ	ʔam	/m/
◇	◦	ब	B b	बु	bu	/b/
U	ʋ	प	P p	पु:उ	puʔ	/p/
U	ᷗ	ह	H h	ह्ओ	hiyɔ	/h/, Vokaldehnung
℧	ᴣ	न	L l	ह्ओ	hɔlɔ	/l/
℥	5	ड	Ṛ ṛ	हड	hɔr	/r/
T	ℽ	र	R r	हर्	har	/r/
ℍ	℥	ष,श	Ṣ ṣ	षु	śu:	/ś/, /s/
ʮ	℀	स	S s	सो	si:	/s/
℧◇ℴℴ		ह्ब	HB hb	—	—	/b/, /ɯ/

5. Zeichen für die heilige Silbe ℗ OM

6. Zahlzeichen

ℐ 1	ℰ 2	ℳ 3	५ 4	ℒ 5	ℳ 6	℮ 7	ℐ 8	ℂ 9
℧ 10	✕ 20	✕ 30	ℒ 40	ℒ 50	ℳ 60	℮ 70	ℎ 80	ℂ 90
aber ℐℐ 11	ℰℳ 23	usw.			ℒℰ 57			

Ɛⴸᴧ ०२४३ – स्वर बर्णं – Vowels

ᄓ₆₆ ९६ – अक्षरे' – Alphabets

४ⴸ ८७ बड़ा	ⴤᴌⴑ₣₃ ८७ छोटा	४ᴌ₅ᴌ₃ उच्चारण	ⴑᴧ₣ᴧ♈₣ हिन्दी	Ɛ₃ⴝ₆₣₣४ अ ग्रेजी
Capitals	**Small**	**Pronunciation**	Hindi	*English*
V	ν	ङ॒	॑	n̂
ʃ	ʃ	s, अ॒	T, अ	a
ꟓ	₣	वि:ड़	ि	i
L	ι	शु:उ	ꞈ	u
Y	y	श:s	:	?
ᴋ	ᴘ	यो	य	eə
Ɛ	₴	इ	ई, ी	ii
ⴡ	ꓷ	उ	उ, ॖ	uu
ⴸ	ꓷ	स꘎	॔, ए	e
Z	ꓜ	꘎अ	ो, औ	o

◇ⴸ₃ᴧ ०२४३ – व्यञ्जन वर्णं – Consonant

ⴝ	ꝫ	꘎ङ॒	ङ	ng
ⴖ	ꓵ	ग्॒	ग्॒	g
�9	ꝫ	क्॒	क्॒	k, ch

Table 7. The Ho alphabet
(*Source*: *Bah Buru Bonga' Buru,* an intermediate Ho reader; title in Hindi:
Hoo Padya-Gadya Saṃgraha [Ho Verse-Prose Collection], Ranchi, 1984.)
(*Information parallel to but not identical with that in Table 6.*)

(Table 7, continued)

				transliteration
ॐ	ೲ	विज्	ज्	n̠
E	३	विज्	ज्	j
प	७	विच्	च्	ch, c
ॻ	ℵ	एण	ण्	N
ा	ſ	ऐड्	ड्	d
L	ι	टे:ए	ट्	t
↿	τ	नुङ	न्	n
ॺ	⌣	द:s	द्	d.
⊖	६	तेत्	व्	t.
♢	৪	sझ	म्	m
♢	०	बू	ब्	b
∪	∪	पु:उ	प्	p
℧	ᔕ	ह्यो	ह	h
℧	℘	ह्ल	ल्	l
5	६	हड़	ड्	l., r
ॻ	⤬	हs	र्	r
Ⴗ	⅄	षू	ष्	sh
⅄	৮	सी	स्	s

The aspirates are not found in Ho words, and are not written in Deeney. In VK, in Table 8 on the Ho numerals, where a transcription of the Hindi (Devan⁻agarī) numerals into VK is given, consonantal aspiration (in writing the number 'six') is indicated by the {h} character used as a superscript diacritic, written above the consonant. The {h} is used—bafflingly—in Hindi 'four', where Hindi (and Devanāgarī) have no aspiration. The {h} here is written after the initial consonant, contrasting with the "vowel lengthening" which is written not after the consonant but after the vowel.

As to how preglottalized stops are handled, see the discussion of this for SS and OC. The two preglottalized stops found in Ho, /ʼb/ and /ʼḍ/, are unmarked: in final position they are realized as preglottalized stops; word medially before a vowel they are read as voiced stops (with no preglottalization). This probably leads to some ambiguity in word-final (and some word-medial syllable-final) occurrences in recent borrowings where the /b/ and /ḍ/ are not preglottalized and are not distinctively written. It seems that they are also unmarked in Deeney's Devanāgarī Ho. Deeney often indicates in his glossary entries where final {ḍ} is preglottalized (in his Devanāgarī explications, not in the Roman main entry listings). He does this with a diacritic, but a similar disambiguation for {b} (rarely needed for either consonant) is made very rarely, if at all. It is only the rare borrowing that would not be preglottalized. As Pinnow notes, another source of ambiguity results from the absence of any *halanta*-like inherent-vowel-deleting diacritic, such as is found in Hindi and other Indo-Aryan languages. Here, the script does not indicate whether there is a short /a/ or, less frequently, an /o/ or /e/, or no vowel at all in certain word-medial consonant sequences.

4.2. On the mysterious {HB}

The {hb} sequence, exemplified in Pinnow's {hbar(a)ŋ} and {hbonga}, presumably represents a distinct pronunciation in HoH, a distinct phoneme (Pinnow /b/, /v/, the latter also transliterated in VK publications as {w}). It is not clear whether the (hb) is sometimes pronounced /w/ and sometimes /b/ (indistinguishable from {b}?), and if so, when. The cognate forms, insofar as I can identify them (e.g., **boŋga** 'god, spirit', for VK **hboŋga**) have (or have merged to?) /b/, presumably with no phonetic difference observed by Deeney for HoK. It is possible that words with {hb} are semantically marked as somehow 'hieratic', in addition to any (independent or not) phonetic distinction.

It is not at all clear what is distinctive about HoH—some phonetic distinctions, or perhaps a graphemic distinction (assuming that HoK is written or can be conceived of as written)? Certainly there are lexical differences. In the latter how much can be found in the way of different lexemes, and how

Norman Zide

᧬	ᨩ	ᨰ	ᨯ	ᨳ
ᨳᨴᨯᨱ	ᨻᨴᨯᨱ	ᨶᨻᨳᨱ	ᨶᨭᨲ	ᨳᨴᨱ
मोयङ्	ऋइया	पिया	पुन	मोङ्
१	२	३	४	५
एक	दो	तीन	चार	पांच
ᨯᨴᨮ	ᨳᨰ	ᨱᨦᨴᨲ	ᨲᨴᨱᨳ	ᨶᨱᨶᨲ

ᨴ	ᨳ	ᨳ	ᨳ	ᨳ
ᨱᨳᨳ	ᨳᨴᨯᨱ	ᨶᨻᨴᨲ	ᨳᨮᨴᨯᨱ	ᨴᨮᨮᨶ
तुर	इया	रिल्ल	रेया	गेल
६	७	८	९	१०
छव	सात	आठ	नव्	दस
ᨳᨱᨳ	ᨳᨱᨳ	ᨴᨱᨲ	ᨴᨳ

Table 8. *The Ho numerals from 1 to 10. Ho forms in VK and Devanāgarī, and Hindi forms in Devanāgarī and VK.*
(Source: Deeney 1975.)

(Table 8, continued)

moyad (VK)	riya	piya	pun	mod
moya'd (Dev)	ryā	piyā	pun	mo'd
1	2	3	4	5
eHk (VK)	do	tiHn	char	paṃc
ek (Dev)	do	tīn	cār	pāc

****

tur	iya	ril	reya	gel
tur	iyā	tīn	reyā	gel
6	7	8	9	10
čao	sat	yat	nao	d()s
chav	sāt	āth	nav	das

Note: The '*d* represents a preglottalized (retroflex) *d*. The *r* in *ryā* 'two' represents the Sanskrit and Devanāgarī character for the Sanskrit vowel *r* . The tilde in *pāc* 'five' represents the *candrabindu* vowel nasalization, but in *čao* 'six' the tilde represents the VK {h} used as a diacritic (superscript) for aspiration of consonants, rarely, and probably always in clearly identified loans or a transcription of Hindi. () represents the "inherent vowel", here "short *a*" or schwa. The {H} represents vocalic lengthening.

much special interpretation is there of common (to HoH and HoK) vocabulary?
Pinnow suggests that the HB is old in Austroasiatic although he finds no
cognate phenomena with it in the Munda languages, and derives it from a
preglottalized *'b* (as in Mon-Khmer and presumably Proto-Austroasiatic). He
relates it to prenasalized consonants in Mon-Khmer, and suggests that
{hboŋga} may be cognate with very distantly related Lamet /mbrooŋ/ '*Geist*;
Dämon'. The present author suspects that is is more likely the result of a recent
dialect split (within Kherwarian, if not Ho), the HoH ('shamanic dialect'?)
making the previously allophonic HB (i.e., *w* vs. *b*) distinction phonemic, a
distinction that HK lacks, and according to Pinnow must have lost.

For the /y/ and /ʔ/, see the discussion of "ligatures" below (following
Pinnow following Lako Bodra). Pinnow points out that the VK sign for OM is
occasionally used as a ligature sign, e.g., in {ṣOMs} corresponding to Hindi
śvāṃs 'breathing'.

4.3. Vowels, mixed vowels and ligatures

This follows Pinnow's presentation of the VK material closely. VK has ten
vowel characters, broadly labeled "simple vowels," "mixed vowels," and
"ligatures." (See again Tables 5 and 6.) VK has four "simple vowels" (*īpan
bor(o)ŋ*; plus the "inherent vowel", the latter feature (and the notion of /e/ and
/o/ as not simple) being held in common with the Devanāgarī and other Indic
scripts. The non-inherent simple vowels are /a/, /i/, /u/, and what Pinnow
transcribes as /m̥/, /ŋ/. This last "simple vowel" is the vowel nasalization
character corresponding to the *candrabindu* and *anusvāra* diacritics in Hindi. As
Pinnow's table shows, the inherent vowel is not always to be pronounced /a/,
but sometimes /o/, and more occasionally /e/; it is the usual representation of /a/,
and represents /o/ and /e/ (everywhere ? consistently?) where they are
"automatic," i.e., the second vowel in a dissyllabic word where the second is
identical with the first.

The two "mixed vowels" are /e/ and /o/; this presumably follows Sanskrit
notions, according to which these are not "simple." (What Lako Bodra means
by *īpan bor(o)ŋ*, and how more precisely /e/ and /o/ are mixed, is not clear to
me, though it is probably more or less as in Sanskrit /e/ = /a+i/. The *dobṛi
bor(o)ŋ* (Pinnow's VK "ligatures"), neither vowels (simple or mixed) nor
consonants, are four (see Table 6): the sign representing /ʔ/, the sign for /y/ (and
variants), and the characters for 'long i' and 'long u' (presumably these two are
long, and thus not simple, but also not mixed, i.e., combined of different
elements). The glottal stop /ʔ/ is unequivocally represented by a character that
seems to be based on the /m̥, ŋ/ character. (The transliteration into Devanāgarī
gives the *visarga* sign.)

The characters for the "ligatures" are themselves composite; they seem to correspond with, or be built on, the four simple vowels. I can see why the long /ū/ and /ī/ characters are most likely built from the short vowel characters, but how the /ʔ/ and /y/ relate to /m̩, ŋ/ and /a/ is far from clear. Or is the seeming matching an arbitrarily imposed pictorial neatness, with no phonetic implications? The {y} character, according to Pinnow, can represent /y/, /e/, or /i/; the details of that will not be discussed here. I do not take the character shapes for /e/ and /o/ to be derived from any of the simple vowels, although a (weak) case for that can be made. Pinnow writes that Devanāgarī {ai} and {au} are represented in VK by {ī} and {ū}.

4.4. *Ho Hayam vs. Ho Kaji*

Below is a sample lexical set, consisting of the numbers from one to ten. (Pinnow also discusses the number words.)

Ho Hayam
moya'd '1', **riya** '2', **piya** '3', **pun** '4', **mo'd** '5', **tur** '6', **iya** '7', **ril** '8', **reya** '9', **gel** '10' (from *Ho Hayam Paham Puti*)

Ho Kaji (from Deeney; Pinnow's forms are very slightly different; both full and short forms are given for each number)
miya'ḍ/mi'ḍ '1', **bariya/bar** '2', **apiya/apē** '3', **upuniya/upun** '4', **mõya/moe** '5', **turuiya/turui** '6', **aiya/ai** '7', **irleya/iril** '8', **areya/ arē** '9', **geleya/gel** '10'

A fairly regular truncation pattern can be seen deriving the HoH forms from something very like the modern HoK forms, with interseries assimilation increased, e.g., '2' and '3' (**riya** and **piya** in HoH), and '2' and '9' (**riya** and **reya**). All the HoH numerals except '1' are either monosyllabic (ending in a non-grave consonant: either the sonorants *r, l, n* or the preglottalized stop *'d*) or end with /ya/, which elsewhere is a distinct morpheme; here presumably the numerals are not syntactically distinguished between those with the morpheme *ya* and those lacking it). There are a few unexpected features, e.g., the final preglottalized /'d/ in '5', which may owe something to assimilation to the form for '1'. Note the assimilation of 'two' /riyā/ (a peculiar VK form) to a truncated (with initial syllable, i.e., V-, deleted) analogue of standard 'three' /piyā/. '9' in HoH also loses its initial syllable. The HoH forms, on the whole, do not look archaic.

The VK transcriptions of the Hindi forms of the numbers in Table 8 show some unexpected handling of Hindi vowels and aspiration, which conceivably

could suggest that Lako Bodra heard in the sounds of Ho and perceived in the making of those sounds (in his efforts to dissimilate the script from Devanāgarī) a different phonation type from Hindi. Mundari has a distinctive phonation type which has not been described; it is not, however, contrastive with any other phonation type within the language, and may instead be some different form of laryngeal modification. No one has yet done any myographic or other studies on Mundari speakers.

Historically—considering the almost vestigial tone/register distinction in (North Munda) Korku and the complex system (e.g., four way: plain, breathy, creaky and creaky plus breathy voice) of (Mon-Khmer) Cong—it is quite possible that Proto-North Munda had distinctive phonation types, and that Ho phonetics (Ho Hayam phonetics as opposed to Ho Kaji?) retains some of the properties of one "non-plain" phonation type. The pecularities in the transcription of Hindi (someone's standard Hindi) in the VK numbers are the following:

•the 'lengthened' vowel in '1' and '3' for the (long) front vowels /e/ and /i/,
•aspiration (of some kind, contrasting with the lengthening just mentioned and with the "regular" [?] consonantal aspiration, /ch/ for '6').
•Are the *e* and *ī* of Hindi /ek/ and /tīn/ transcribed differently (i.e., {eH} and {iH}) because they are heard—even if not consistently—as different, or is this just 'careless' hit-or-miss transcription? Note also the unexpected transcription—not transliteration—of Hindi /āṭh/ '8', VK /yaṭ/. The loss of the final aspiration is not surprising, but the /y/ + short /a/ is. The complications with '6' and '9' have to do with final consonants or semivowels, and there are several variant forms in Hindi for these numbers, so we will not speculate about the Ho forms here. There is, of course, the problem of VK transliteration from Ho versus transcription of someone's authoritative pronunciation, and perhaps a mixture of the two in the VK versions of the Devanāgarī forms given above, but I do not think this makes the data less interesting or significant.

REFERENCES

BIBLE SOCIETY OF INDIA AND CEYLON. n.d. *Proverbs in Ssaora (baŋsa ənobleganji).*

_____. 1961. *Mattiu (Matthew).*

BODDING, P. O. 1923. *Materials for a Santali Grammar.* Part 1. Benegaria: Santal Mission Press.

BURROWS, L. 1915. *Ho Grammar (with Vocabulary).* Calcutta: Catholic Orphan Press. (Reprinted as *The Grammar of the Ho Language: An Eastern Himalayan Dialect,* New Delhi: Cosmo, 1980.)

DEENEY, J. 1975. *Ho Grammar and Vocabulary.* Chaibasa: Ho Xavier Publications.

_____. 1978. *Ho-English Dictionary.* Chaibasa: Ho Xavier Publications.

LAKO BODRA. 1963. *Ho Hayam Paham Puti.* Jhinkpani. (Also several other pamphlets listed by Pinnow, which I have not seen.)

_____. (n.d.) *Ho Halaŋ Galaŋ.* Jhinkpani. (A trilingual dictionary, the manuscript of which is in the possession of Mr. B. Pat Pingua of Ranchi, a follower of Lako Bodra, who also provided information on Lako Bodra's movement and the script, information sometimes at variance with that of Pinnow's informants.)

MAHAPATRA, B. P., and Ranganayaki MAHAPATRA. 1979. "Santal script and texts." Working paper for the post-plenary session of the Tenth International Congress of Anthropological and Ethnological Sciences, Mysore), 1-15.

MAHAPATRA, Kh. 1978-79. "'SoraN SompeN': A Sora script." Unpublished conference paper (Delhi, Mysore). (Oriya version in *oria lipi o bhāṣā,* Cuttack, 1977, pp. 147-155.)

PINNOW, H.-J.. 1972. "Schrift und Sprache in den Werken Lako Bodras im Gebiet der Ho von Singbhum (Bihar). *Anthropos* 67:822-855.

SMALLEY, William A., Chia Kona VANG, and Gnia Yee YANG. 1990. *Mother of Writing: the Origin and Development of a Hmong Messianic Script.* Chicago and London: University of Chicago Press.

ZIDE, N. H. 1967. "The Santali Ol Cemet' script." In *Languages and Areas: Studies Presented to George V. Bobrinskoy* [Chicago: privately printed], 180-189.

_____. 1968. "Graphemic system in the Ol Cemed script." In *Papers from the Fourth Regional Meeting of the Chicago Linguistic Society*, 238-254.

_____. 1991. "Lexicography of Other Languages of the Indian Subcontinent: The Munda Languages". In *Wörterbucher: Ein internationales Handbuch zur Lexicographie*, ed. by F. J. Hausmann, O. Reichmann, H. E. Wiegand, and L. Zgusta, 2533-2547. Berlin and New York: Walter de Gruyter.